

Artificial Intelligence Bias Audits Policy Template



Savvy use of software, algorithms, and artificial intelligence (which we'll refer to collectively as "AI") can help you take your hiring process to the next level. But it can also get your company into serious legal trouble. The same digital technology that makes it possible to identify and screen job candidates can also be used, whether deliberately or inadvertently, to exclude disadvantaged groups protected by human rights laws. One way to protect yourself is to perform what are called "AI bias audits" to determine whether your AI hiring applications may be inappropriately skewing against protected groups. Here's a template AI Bias Audits Policy that you can adapt for your own use.

1. POLICY

While reserving the right to develop, obtain, and use artificial intelligence (AI) tools (as those terms are defined below) to increase the efficiency and effectiveness of its recruitment, hiring and other HR functions, ABC Company acknowledges that AI tools used for these purposes may have or develop bias that can subject job applicants to discrimination while exposing the Company to risk of liability under human rights and other applicable laws. Accordingly, ABC Company will ensure that bias audits of the AI tools it uses are carried out periodically for purposes of preventing employment discrimination through algorithmic bias in accordance with this Policy.

2. DEFINITIONS

For purposes of this Policy:

"Artificial Intelligence" (AI) refers collectively to a broad range of computer, algorithmic, and computer decision-making tools used to assist in making hiring and other employment decisions, including but not limited to software for video interviewing, resume-screening, and other hiring applications (or "apps"), algorithms, algorithmic decision-making tools, machine learning, machine-based systems that can make predictions, recommendations or decisions influencing hiring, and other employment decisions and chatbots;

“AI tool” includes products that use artificial intelligence, machine learning, statistical modeling, data analytics, or mathematical, computer-based techniques: i. to generate either a prediction, such as an assessment of a job applicant’s fit or likelihood of success, or a classification, that is, an assignment of an observation to a group, such as categorizations based on skill sets or aptitude; and ii. for which a computer at least in part identifies the inputs, the relative importance placed on those inputs, and, if applicable, other parameters for the models to improve the accuracy of the prediction or classification;

“Category” means race, sex, or other personal characteristics protected from discrimination or that can be used as a comparator against groups possessing such a personal characteristic for purposes of determining whether discrimination occurred, e.g., males as a comparator to females;

“Historical data” means data collected during use of an AI tool to assess job applicants for employment;

“Impact ratio” means either: (1) the selection rate for a category divided by the selection rate of the most selected category; or (2) the scoring rate for a category divided by the scoring rate for the highest scoring category;

“Scoring Rate” means the rate at which individuals in a category receive a score above the sample’s median score, where the score has been calculated by an AI tool;

“Selection rate” means the rate at which individuals in a category are either selected to move forward in the hiring process or assigned a classification by an AI tool, which may be calculated by dividing the number of individuals in the category moving forward or assigned a classification by the total number of individuals in the category who applied for a position.

3. SCOPE OF POLICY

This Policy applies to ABC Company’s use of AI tools for performing recruiting and hiring functions. In the event that ABC Company expands its use of AI tools for other employment purposes, such as assessing whether to promote current employees, this Policy will be expanded to cover those additional functions.

4. SELECTION & DEPLOYMENT OF NEW AI TOOLS

ABC Company will exercise due diligence in selecting AI tools before using them by evaluating products and vendors in accordance with standards established by human rights commissioners, privacy commissioners, and other regulatory authorities, as well as ABC Company’s own anti-discrimination and privacy policies and protocols. After determining that an AI tool meets these standards, ABC Company will select an external, independent auditor to review and test the new AI tool for bias. If the audit finds that the AI tool does not have algorithmic bias, ABC Company will initiate parallel testing using the AI tool to perform current practices, e.g., HR will screen resumes alongside the AI tool and compare their respective outcomes) for [amount of time] to verify whether bias exists in practice.

ABC Company will deploy the new AI tool only after thoroughly vetting the vendor and conducting a bias audit and parallel testing to verify that no bias was found.

5. ANNUAL BIAS AUDITS

Neither ABC Company nor any hiring or employment agency with which it contracts to perform or assist in hiring functions may use or continue to use an AI tool if more than one year has passed since the most recent bias audit of that tool. AI bias audits will be performed by an external, independent auditor selected by ABC Company.

5.1 Bias Audits of Classification AI Tools

Where an AI tool selects candidates for employment to move forward in the hiring process or classifies them into groups, the annual bias audit must, at a minimum:

- Calculate the selection rate for each category;
- Calculate the impact ratio for each category;
- Ensure that both of the above calculations separately calculate the impact of the AI tool on groups protected from discrimination, including [*list protected groups, e.g., race, sex, national origin, etc.*], as well as intersectional categories of the above;
- Ensure that the above calculations are performed for each group, if an AI tool classifies candidates for employment into specified groups (e.g., leadership styles); and
- Indicate the number of individuals the AI tool assessed that are not included in the required calculations because they fall within an unknown category.

5.2 Bias Audit of Classification AI Tools

Where an AI tool assigns scores to candidates for employment, the annual bias audit must, at a minimum:

- Calculate the median score for the full sample of applicants;
- Calculate the scoring rate for individuals in each category;
- Calculate the impact ratio for each category;
- Ensure that all of the above calculations separately calculate the impact of the AI tool on groups protected from discrimination, including [*list protected groups, e.g., race, sex, national origin, etc.*], as well as intersectional categories of the above; and
- Indicate the number of individuals the AI tool assessed that are not included in the required calculations because they fall within an unknown category.

6. HISTORICAL DATA

Annual bias audits must use historical data of the AI tool, which may come from ABC Company or one employer or employment agency that uses the AI tool.

7. MITIGATION OF DETECTED AI BIAS

Where the annual AI bias audit finds substantial evidence of algorithmic bias that caused or likely caused the AI tool to discriminate against protected classes of job applicants, HR/IT/Management must immediately notify the vendor that provided the AI tool and work together with the vendor to establish and implement a plan to remedy past violations and effect corrective actions to ensure that further discrimination

does not occur in the future. If an immediate and effective remedy is not available, ABC Company must immediately stop using the AI tool. In this situation, ABC Company may seek a new AI tool to replace the discontinued tool.

8. EMPLOYEE NOTIFICATION OF POTENTIAL AI BIAS

Employees who observe or have reasonable grounds to suspect that an AI tool has developed or could potentially develop bias must immediately alert their supervisor or HR. Upon receiving such a report, HR will perform an internal audit of related data to determine if further investigation by an independent auditor is required. If so, HR or management will contact an independent auditor to arrange an audit.

9. REASONABLE ACCOMMODATIONS

HR will train system administrators on how to identify and handle reasonable accommodation requests from job applicants who have disabilities that might prevent them from using the AI tool effectively. ABC Company will make reasonable accommodations for these applicants up to the point of undue hardship on a case-by-case basis in accordance with the Company's Reasonable Accommodation policy and procedures.